

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-203337

(43) 公開日 平成7年(1995)8月4日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 N 5/76	B			
G 0 6 F 3/06	5 4 0			

審査請求 有 請求項の数18 O L (全 12 頁)

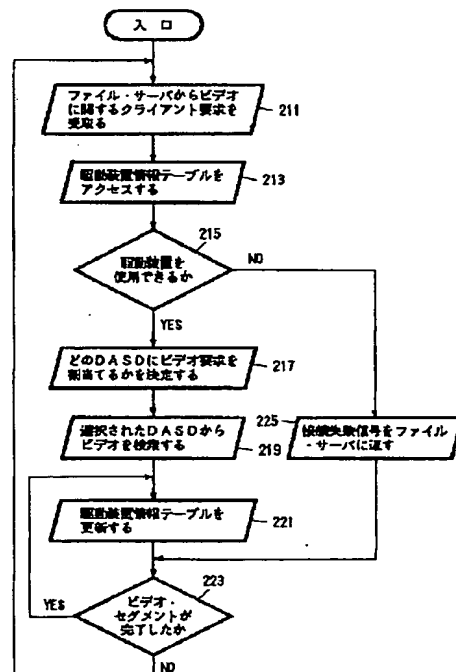
(21) 出願番号	特願平6-288128	(71) 出願人	390009531 インターナショナル・ビジネス・マシー ズ・コーポレーション INTERNATIONAL BUSIN ESS MACHINES CORPO RATION アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)
(22) 出願日	平成6年(1994)11月22日	(72) 発明者	ウェンディ・リン・ヌスピッケル アメリカ合衆国33483 フロリダ州デルレ イ・ビーチ アイリス・ドライブ 908
(31) 優先権主張番号	1 5 9 6 4 7	(74) 代理人	弁理士 合田 潔 (外2名)
(32) 優先日	1993年11月30日		
(33) 優先権主張国	米国 (US)		

(54) 【発明の名称】 RAIDを使用する、フル・モーション・ビデオのネットワーク・サポートの方法および装置

(57) 【要約】

【目的】 RAID型大容量記憶システムからネットワ  
ークの複数のノードで非同期フル・モーション・ビデオ  
を再生するためのシステムおよび方法を提供する。

【構成】 複数の直接アクセス記憶装置のそれぞれに複  
製ビデオ・セグメントを記憶することによって、別個の  
フル・モーション・ビデオ・セグメントを、複数の再生  
プラットフォーム上で再生できる。ビデオ・セグメント  
に関する要求に応答して、ビデオ・セグメントの検索の  
ための直接アクセス記憶装置を、駆動装置情報テーブル  
にリストされた装置の中から選択する。次に、選択され  
た直接アクセス記憶装置に、ビデオ・セグメントを検索  
するよう指令する。最後に、駆動装置情報テーブルを更  
新して、検索に関する選択された直接アクセス記憶装置  
の使用を反映させる。



## 【特許請求の範囲】

【請求項 1】サーバによって制御される複数の直接アクセス記憶装置のそれぞれに、複製ビデオ・セグメントを記憶するステップと、

再生プラットフォームからの要求に応答して、駆動装置情報テーブルからビデオ・セグメントの検索のために直接アクセス記憶装置を選択するステップと、

選択された直接アクセス記憶装置に、ビデオ・セグメントを検索するよう指令するステップと、

検索に関する選択された直接アクセス記憶装置の使用を反映するため、駆動装置情報テーブルを更新するステップと、

ネットワークを介して再生プラットフォームにビデオ・セグメントを伝送するステップとを含む、ネットワークを介してファイル・サーバに接続された複数の再生プラットフォームに、フル・モーション・ビデオを配布する方法。

【請求項 2】前記選択するステップが、複数の直接アクセス記憶装置のそれぞれに関する使用可能性情報を検索するため、駆動装置情報テーブルをアクセスするステップと、

1 つまたは複数の直接アクセス記憶装置が使用可能である場合に、ビデオ要求をサポートするため使用可能な直接アクセス記憶装置のうちの 1 つを割り当てるステップとを含むことを特徴とする、請求項 1 に記載の方法。

【請求項 3】前記選択するステップが、さらに、使用可能な直接アクセス記憶装置がない場合に、再生プラットフォームに接続失敗を知らせるステップを含むことを特徴とする、請求項 2 に記載の方法。

【請求項 4】ビデオ・セグメントに関する再生プラットフォームからの要求に応答して、複数の直接アクセス記憶装置へのアクセスを制御するルータに要求を転送するステップをさらに含む、請求項 3 に記載の方法。

【請求項 5】前記選択するステップが、ルータから実行されることを特徴とする、請求項 4 に記載の方法。

【請求項 6】前記選択するステップが、装置ドライバから実行されることを特徴とする、請求項 3 に記載の方法。

【請求項 7】必要データ速度を、拡張属性としてビデオ・ファイルに付加するステップをさらに含む、請求項 3 に記載の方法。

【請求項 8】複数のビデオ・セグメントの再生が、同時にサポートされることを特徴とする、請求項 7 に記載の方法。

【請求項 9】複数のビデオ・セグメントの再生が、複数の再生プラットフォームで発生することを特徴とする、請求項 8 に記載の方法。

【請求項 10】複数の再生プラットフォームと、ファイル・サーバと、

複数のプラットフォームとファイル・サーバとを相互接

続するネットワークと、

複数の直接アクセス記憶装置と、

直接アクセス記憶装置のそれぞれの状況情報を提供する駆動装置情報テーブルと、

複数の直接アクセス記憶装置のそれぞれに記録された、複数の複製ビデオ・セグメントと、

再生プラットフォームからファイル・サーバへのビデオ・セグメントに関する要求に応答して、駆動装置情報テーブルへの参照によってビデオ・セグメントの検索のために直接アクセス記憶装置を選択する手段と、

選択された直接アクセス記憶装置に、ビデオ・セグメントを検索するよう指令する手段と、

検索に関する選択された直接アクセス記憶装置の使用を反映するため、駆動装置情報テーブルを更新する手段とを含む、データ処理システム。

【請求項 11】前記選択する手段が、複数の直接アクセス記憶装置のそれぞれに関する使用可能性情報を検索するため、駆動装置情報テーブルをアクセスする手段と、

1 つまたは複数の直接アクセス記憶装置が使用可能であることに応答して、ビデオ要求をサポートするため使用可能な直接アクセス記憶装置のうちの 1 つを割り当てる手段とを含むことを特徴とする、請求項 10 に記載のデータ処理システム。

【請求項 12】前記選択する手段が、さらに、使用可能な直接アクセス記憶装置がないことに応答して、再生プラットフォームに接続失敗を知らせる手段を含むことを特徴とする、請求項 11 に記載のデータ処理システム。

【請求項 13】ビデオ・セグメントに関する再生プラットフォームからの要求に応答して、複数の直接アクセス記憶装置へのアクセスを制御するルータに要求を転送する手段をさらに含む、請求項 12 に記載のデータ処理システム。

【請求項 14】前記選択する手段が、ルータであることを特徴とする、請求項 13 に記載のデータ処理システム。

【請求項 15】前記選択する手段が、装置ドライバであることを特徴とする、請求項 12 に記載のデータ処理システム。

【請求項 16】必要データ速度を示す、選択されたビデオ・ファイルに対する拡張属性をさらに含む、請求項 12 に記載のデータ処理システム。

【請求項 17】複数のビデオ・セグメントの再生が、同時にサポートされることを特徴とする、請求項 16 に記載のデータ処理システム。

【請求項 18】複数のビデオ・セグメントの再生が、複数の再生プラットフォームで発生することを特徴とする、請求項 17 に記載のデータ処理システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、ネットワークの複数のノード上でのRAID型大容量記憶システムからの非同期フル・モーション・ビデオの再生のシステムおよび方法に関する。

#### 【0002】

【従来の技術】ディスク記憶装置は、不揮発性であり、メモリ・サイズの需要が実用的な主記憶装置の量を追い越し続けているので、コンピュータにおいてディスク記憶装置の使用は重要である。通常、データ処理アプリケーションの場合、単一のディスクが、システムの中央演算処理装置(CPU)がデータを利用できる速度より遅い速度でデータを供給する。主記憶装置は、ディスクからCPUへの緩衝記憶装置として働くが、通常のアプリケーションのシステム性能は、ディスクのアクセス速度によって制限されることがしばしばである。したがって、総合システム性能を向上させるには、ディスク駆動装置の記憶サイズとデータ・アクセス速度を高めることが必要であると考えられている。これに関しては、Michelle Y. Kim著、"Synchronized Disk Interleaving", IEEE Transactions On Computers, Vol. C-35, No. 11, November 1986を参照されたい。

【0003】データ・アクセス速度を高めるために、さまざまな技法が使用されてきた。1トラック全体のデータを保持できるディスク・キャッシュ・メモリが、単一トラック上のデータの連続アクセスの場合のシーク遅延と回転遅延を除去するのに使用されている。複数の読み書きヘッドが、単一ディスクのトラックの組または1組のディスク上でデータのブロックをインターリーブする(データ・ストライピング)のに使用されている。一般的なデータ・ブロック・サイズは、バイト・サイズ、ワード・サイズおよびセクタ・サイズである。ディスク・インターリーブは、性能向上のための既知のスーパーコンピュータ技法であり、たとえば、上に記載の論文で論じられている。

【0004】データ・ストライピングは、データ帯域幅に関しては有利であるが、残念ながら、故障までの平均時間が短くなる。これは、データを記憶するのに使用されるアレイ内のディスクの台数が増えるにつれて減少する。この、システムの故障までの平均時間が短くなるのを補正するために、エラー認識とエラー訂正をデータに追加して、いわゆるRAID (Redundant Arrays of Inexpensive Disks) アーキテクチャが作られた。RAIDアーキテクチャの5つのタイプを、RAIDレベル1ないしレベル5と称する。David A. Patterson他著、"A Case for Redundant Arrays of Inexpensive Disks (RAID)", Report No. UCB/CSD 87/89, December, 1987, Computer Science Division (EECS), University of California, Berkeley, California 94720を参照されたい。

【0005】RAIDレベル1では、主ディスクからバックアップ・ディスクへのデータの完全な複製(ミラー

リング)を利用する。RAIDレベル1は、主に、耐故障性をもたらすためのデータ冗長方式とみなされてきた。通常の駆動装置に対してディスクあたりの性能比が相対的に低いので、RAIDレベル1は、他の分野ではほとんど使用されていない。

【0006】RAIDレベル2は、RAIDレベル1の性能を高め、データおよびディスクの障害回復を提供するのに必要な余分なディスクの数を減らすことのできる誤り訂正コードを利用することによって、ディスクあたりの容量比を高めたものである。RAIDレベル2では、データをG枚のディスクのグループにインターリーブし、誤り訂正コード(ECC)を生成し、「チェック・ディスク」と称するC枚のディスクの追加セットに記憶して、単一のエラーを検出し、訂正する。ECCは、データ内のランダムな単一ビットのエラーを検出し、訂正できるようにすると同時に、G枚のデータ・ディスクのうちの1枚が破損した場合にデータを回復できるようにするのに使用される。C+G枚のディスクのうちのG枚だけがユーザ・データを保持するので、ディスクあたりの性能は、 $G/(G+C)$ に比例する。 $G/C$ は、通常は1よりかなり大きく、したがって、RAIDレベル2は、RAIDレベル1に対してディスクあたりの性能が向上している。1つまたは複数の予備ディスクをシステムに含めることができ、その結果、ディスク駆動装置のうちの1つが故障した場合に、予備ディスクに電子的に切り替えて、RAIDシステム内の故障したディスクと交換できるようになっている。

【0007】RAIDレベル3は、RAIDレベル2の変形であり、既存のほとんどの安価なディスク駆動装置が提供するエラー検出能力を利用して、チェック・ディスクの数を1つに減らし、これによってRAIDレベル2に対してディスクあたりの相対性能を向上させたものである。通常、パリティ・データが、ECCの代わりに使用される。ECCその他のエラー・コードまたはパリティ・データのいずれかを、「冗長データ」と称する場合がある。RAIDレベル2およびレベル3のどちらの場合でも、データ・ディスクの全てへの帯域幅を利用できるので、大量のデータまたはグループ化されたデータのディスク・アクセスのトランザクション時間が短縮される。

【0008】トランザクション処理に一般的であるものなどの少量のデータ転送に関する性能判断基準は、RAIDレベル2およびレベル3では低いことがわかっている。というのは、データが、ビット・サイズのブロックでディスクにまたがってインターリーブされており、1セクタ未満のデータのデータ・アクセスであっても、すべてのディスクをアクセスしなければならないからである。この性能パラメータを向上させるため、RAIDレベル3の変形であるRAIDレベル4では、データが、レベル1ないしレベル3のようにビット・インターリーブ

・モードではなく、セクタ・インタリーブ・モードでディスク上でインタリーブされる。言い換えると、個々の入出力転送に、1つのデータ・ディスクだけが使用される。この利点は、入出力動作の並列性の可能性から生じるものである。これによって、同時に同一のデータ・ディスクをアクセスする、別々のデータ・アクセス要求の間の競合の量が減る。

【0009】RAIDレベル4の性能は、書き込み動作中のチェック・ディスクに関するアクセス競合があるもので、制限されたままである。すべての書き込み動作に関して、データが書き込まれるデータのストライプ（すなわちセクタの行）のそれぞれについて、チェック・ディスクに更新されたパリティ・データを記憶するために、チェック・ディスクをアクセスしなければならない。Patterson他の観察によれば、RAIDレベル4およびレベル5では、単一セクタへの個々の書き込みが、1つの論理大容量記憶装置のディスクの全てを用いるわけではない。というのは、チェック・ディスクのパリティ・ビットが、グループ内の対応するデータ・ビットの全ての単一の排他的論理和に過ぎないからである。RAIDレベル4では、少量のデータ・アクセスであっても、書き込み動作は互いに干渉する。RAIDレベル4の変形であるRAIDレベル5では、パリティ・チェック・データとユーザ・データをすべてのディスクに分配することによって、書き込み動作時のパリティ・データへのアクセス競合の問題が軽減されている。

【0010】さまざまなRAIDアーキテクチャが、帯域幅の増大に対するデータ処理の需要に答えることに成功してきたが、複数のプロセッサ上での独立なビデオ・セグメントのリアル・タイム再生へのストライプ式ディスクまたはRAIDの応用は、それほど成功していない。フル・モーション・ビデオの再生は、データのバーストではなく高いデータ速度の維持を必要とするという点で、ファイルのコピーやスプレッドシートのアクセスなどの通常のデータ転送と異なる。このデータは、時間にクリティカルでもある。データ配布に0.5秒ほどの遅延があっても、再生されるビデオが崩れる結果となる。ネットワークを介する複数のプラットフォームでのフル・モーション・ビデオの再生では、競合の問題も導入されるので、データ回復が複雑になる。

【0011】RAIDレベル4およびレベル5では、独立のデータ要求元のすべてが、ネットワークを介してまたま1つの駆動装置に要求を集中させる場合がある。この場合、不可避免的に、リアル・タイム再生の必要を満たすのに十分な速さでデータを回復できなくなる。

【0012】ストライピングやRAIDレベル2およびレベル3も、これができない。ストライピングとRAIDレベル2およびレベル3は、大量のデータに関する少数の順次要求を満たす目的のものである。この目的を満たすために、帯域幅を増加させたのである。しかし、デ

ィスクが従来通りに同期化されているので、シーク時間や回転待ちに関しては、性能向上が全くなされていない。複数の非順次要求に対応するという点では、何の利益も得られない。

【0013】さまざまなRAIDアーキテクチャは、入出力性能とCPU性能の間のギャップを狭めるための道具と見なされていた。しかし、マルチメディア・データは、通常は、従来のデータ処理に適さない圧縮フォーマットで記憶される。圧縮によって、ディスクとの帯域幅の問題も軽減される。画像を記憶するのに必要なデータは膨大な量であるから、数分以上のビデオをディジタル形式でディスクに記憶する場合には、ビデオ・データの圧縮が必要条件になる。これらのフォーマットの伸長は、プロセッサ集中の動作である。実際、何らかのタイプの圧縮されたディジタル・ビデオ・データの処理については、Intel社の8088、80286および80836SXマイクロプロセッサは、従来のディスク駆動装置がデータを回復する速度についてゆけない。したがって、ネットワーク上でのビデオ再生の処理という問題は、CPU動作速度とディスク駆動装置のデータ回復速度の間の不一致という従来の問題ではなく、ディスク駆動装置からのデータ・ディスパッチの問題である。完全に対話式のマルチメディア・アプリケーション、特にフル・モーション・ビデオおよびオーディオの、複数のネットワーク・ノードでのサポートまたはマルチタスク式コンピュータで走行する複数のアプリケーションのためのサポートは、従来のデータ処理とは異なる問題を提示する。

【0014】

【発明が解決しようとする課題】本発明の目的は、RAID型大容量記憶システムからネットワークの複数のノードで非同期フル・モーション・ビデオを再生するためのシステムおよび方法を提供することである。

【0015】

【課題を解決するための手段】本発明の上記その他の目的は、下記に従って達成される。本発明は、ビデオ・セグメントの再生を求める独立の要求を行う複数の再生プラットフォームのためにマルチメディアをサポートするRAIDシステムを提供する。本発明の方法によれば、複数の直接アクセス記憶装置のそれぞれに複製ビデオ・セグメントを記憶することによって、複数の再生プラットフォームでフル・モーション・ビデオが再生される。ビデオ・セグメントに関する要求に回答して、ビデオ・セグメントの検索用の直接アクセス記憶装置を、駆動装置情報テーブルにリストされた装置の中から選択する。次に、選択された直接アクセス記憶装置に、ビデオ・セグメントを検索するよう指令する。最後に、駆動装置情報テーブルを更新して、検索に関する選択された直接アクセス記憶装置の使用を反映させる。ビデオ・セグメントが完了した時にも、駆動装置情報テーブルを更新し

て、装置の解放を反映させる。

#### 【0016】

【実施例】ここで図1を参照すると、トークン・リング通信リンク10を介する複数のワークステーション12 Aないし12 C上でのフル・モーション・ビデオ再生をサポートするためのローカル・エリア・ネットワーク8のブロック図が示されている。トークン・リング・ジオメトリが図示されているが、本発明は、ローカル・エリア・ネットワークならびに広域ネットワークの他のジオメトリに適用可能である。しかし、使用されるネットワークは、1つのノードが通信チャネル全体のトラフィックを独占できない均等分散ネットワークであることが好ましい。サーバ13は、ワークステーション12 Aないし12 Cのそれぞれにデータの記憶および回復を提供し、本発明を実行するようにプログラムされた、IBM Personal System/2などの通常のパーソナル・コンピュータまたはIBM RS/6000ミッドレンジ・コンピュータ・システムによってサーバ13を提供することができる。サーバ13には、中央演算処理装置60、メモリ64および、出力伝送をフォーマットし、入力伝送をフォーマット解除するためのネットワーク・アダプタ31が含まれる。サーバ13には、サーバ処理用の記憶域を提供する通常の直接アクセス記憶装置(DASD)56を含めることができる。動作中、メモリ64は、オペレーティング・システムおよびLANサーバを構成するルーチンの組(以下ではオペレーティング・システムと呼称する)66ならびに他のオブジェクト68のための記憶域を提供する。

【0017】フル・モーション・ビデオ・データは、複数の直接アクセス記憶装置(DASD)41 Aないし41 Dに記憶される。フル・モーション・ビデオ・セグメントのそれぞれは、DASD41 Aないし41 Dのそれぞれで完全に複製またはミラーリングされる。すなわち、DASD41 Aないし41 Dと同数のセグメントのコピーがある。DASD41 Aないし41 Dは、RAIDシステム21を構成する。フル・モーション・ビデオ・データ・ファイルは、デジタル・ビデオ・インタラクティブ(DVI)技法、IBM社のUltimotionまたはIntel社のIndeoなどのデジタル・ビデオ技術を使用して記憶される。デジタル・ビデオを用いると、ユーザが、毎秒16MBのトークン・リングLANを介してフル・モーション・ビデオを再生できるようになる。LANの各クライアントは、異なるファイルまたは同一ファイルの異なる部分を再生できる。使用されるデジタル・ビデオ・ファイルは、さまざまな圧縮技法を使用して圧縮される。

【0018】RAIDシステム21は、追加のDASD41 Xを接続することによって拡張できる。RAIDルータ17は、RAIDシステム21にデータ・ブロックを配布し、RAIDシステム21からデータ・ブロック

を回復する。RAIDルータ17は、その名が示す通り、ワークステーション12 Xからのフル・モーション・ビデオ・セグメントの要求を、サービスのため特定のDASDに経路指定する。RAIDルータ17には、内部サーバ・バス58へのインターフェース23が含まれ、この内部サーバ・バス58は、ネットワーク・アダプタ31に接続される。ネットワーク・アダプタ31は、トークン・リング通信リンク10へのインターフェースを提供する。

【0019】構成プロセッサ29が、局所プロセッサ25とRAIDシステム21に接続される。構成プロセッサ29を使用して、新たに接続されたDASD41 Xを動作させて、その装置から維持できるビデオ・データ・ストリームの最大数に対してその装置の等級を定めるのに使用される性能情報を生成する。DASD41 Xの性能等級を生成した後に、その情報を局所プロセッサ25に渡し、局所プロセッサ25は、緩衝記憶装置27に記憶される駆動装置情報テーブル33に1項目を追加する。駆動装置情報テーブル33の固定パラメータは、DASD41 Xまたは不揮発性RAM26に記憶して、電源断による消失から保護することができる。

【0020】局所プロセッサ25は、不揮発性RAM26に記憶されたものとして図示されている監視プログラム28を実行して、要求割振り機能を実行する。データは、RAIDルータ17からRAIDシステム21へ、複数の入出力制御装置37 Aないし37 Dを介して渡される。入出力制御装置37 Aないし37 Dのそれぞれは、それぞれ局所緩衝記憶装置39 Aないし39 Dへのアクセス権を有し、それぞれDASD41 Aないし41 Dを制御する。

【0021】ユーザは、ワークステーション12 Aないし12 Cを介して、RAIDシステム21に記憶されたフル・モーション・ビデオ・セグメントを独立にアクセスできる。ワークステーション12 Bを、詳細に図示する。ワークステーション12 Bは、トークン・リング通信リンク10を介してサーバ13と通信する。ワークステーション12 Bは、概略的にはサーバ13に類似しており、これには、ネットワーク・アダプタ70、表示アダプタ84、ハード・ドライブ装置90、中央演算処理装置(CPU)82およびアドレス可能なメモリ88が含まれる。ワークステーション12 Bの構成要素は、内部データ・バス83を介してデータを転送する。CPU82は、キーボード、マウスまたはその両方を含むことのできる入力周辺機器80を直接制御する。CPU82は、サーバ13から受け取るビデオ・データを伸長し、メモリ88に記憶されるか表示アダプタ84内のビデオ・バッファに記憶されるビデオ・フレーム92をリフレッシュする。表示アダプタ84は、表示装置86を駆動し、表示装置86上に、フル・モーション・ビデオが再生される。PVIビデオの場合、CPUではなく、表示

アダプタ 84 がビデオ・データを伸長する。メモリ 88 には、それ自体のオペレーティング・システムと、ローカル・エリア・ネットワーク 8 上で通信セッションを確立するのに使用されるコマンド構造 94 が含まれる。

【0022】図 2 は、ローカル・エリア・ネットワーク 8 内のサーバ 14 の代替配置を示す図である。サーバ 14 を、上で述べたサーバ 13 と異なる範囲について説明する。サーバ 14 では、オペレーティング・システム 66 が、ネットワーク・サーバとしてサービスするための IBM OS/2 LAN server 3.0 を含む OS/2 バージョン 2. X オペレーティング・システムである。RAID ルータ 17 の機能は、この場合、オペレーティング・システム 66 に関連する RAID 機能フィルタ・アダプタ装置ドライバ (ADD) 69 によって提供される。DASD 56、CD-ROM 43 および、RAID 制御装置 45 によって実施される RAID サブシステム 47 のすべてを、RAID 機能フィルタ ADD 69 を介して実施される RAID システム内の駆動装置として取り扱うことができる。駆動装置情報テーブル 33 は、この場合、オペレーティング・システム 66 内に置かれる。実際には、RAID サブシステム 47 が、カスケード接続された RAID 装置になり、RAID 機能フィルタ ADD 69 を用いて実施される RAID システムには、単にもう 1 つの DASD 装置に見える。RAID 制御装置 45 は、通常の形で緩衝記憶装置 139 を使用する。

【0023】アダプタ装置ドライバは、ハードウェア依存モジュールであり、装置ドライバ階層の最下層の構成要素である。OS/2 オペレーティング・システムでは、ADD と装置マネージャの間のインターフェースが、ADD が状態機械と同然になり、入出力装置とシステム・メモリの間でデータのブロックを移動する責任を負う形で設計されている。フィルタ ADD とは、OS/2 装置マネージャと、装置インターフェースを駆動している ADD との間に挿入することのできるフィルタリング・アルゴリズムである。このようなフィルタ・アルゴリズムは、ADD モデル装置ドライバとしてパッケージ化される。フィルタ ADD は、装置マネージャと装置インターフェース用 ADD の間の、コール・ダウン・パス (call downpath) に挿入される。

【0024】RAID 機能フィルタ ADD 69 は、RAID サブシステム 47 内の駆動装置への要求ルータとして働く。

【0025】図 3 は、システム・メモリまたはルータ緩衝記憶装置のいずれかに記憶されるデータ構造としての駆動装置情報テーブル 33 を示す図である。駆動装置情報テーブル 33 には、RAID システム内の駆動装置ごとの項目 109 が含まれる。各項目 109 は、駆動装置 ID フィールド 101、状況フラグ・フィールド 103、サポートされる最大ストリーム数フィールド 105

および現在サポートされているストリーム数を示す使用中フィールド 107 を含む、少なくとも 4 つのフィールドを有する。駆動装置 ID フィールド 101 には、ドライブ名識別または英数字文字を含めることができる。状況フラグ・フィールド 103 は、駆動装置が使用可能であることを示す 1 または駆動装置がオフラインであることを示す 0 とすることができる。サポートされる最大ストリーム数フィールド 105 は、整数とすることができるが、本発明のいくつかの実施例では、このフィールドをデータ速度とすることができる。使用中フィールド 107 も、通常は整数であり、サポートされる最大ストリーム数フィールド 105 の数が上限である。サポートされる最大ストリーム数フィールド 105 の値がデータ速度の場合、使用中フィールド 107 の値も、データ速度になる。ハイ・パフォーマンス・ファイル・システム (HPFS) を使用する、OS/2 オペレーティング・システムに基づく LAN サーバの場合、デジタル・フル・モーション・ビデオ要求のデータ速度必要条件を、実際のデジタル・フル・モーション・ビデオ・ファイルに付加される拡張属性 (EA) として記憶することができる。この場合の拡張属性は、必要な再生速度を示す。その場合、ネットワーク帯域幅が、要求元ワークステーションに保証される。

【0026】状況フラグ・フィールド 103 を用いると、システムが、RAID アレイ内の駆動装置故障に対処できるようになる。1 駆動装置が動作可能である限り、アレイは「動作中」であり、ユーザからアクセス可能である。駆動装置故障についてのユーザ通知は、既存の RAID レベル 1 実施態様と一貫していなければならない。たとえば、一部の LAN サーバ実施態様、たとえば OS/2 LAN サーバ・バージョン 3.0 では、アレイ内の駆動装置が故障した時に、ユーザが通知または警報を受けることができる。追加駆動装置の追加または故障駆動装置の交換の場合、ユーザがアレイを再フォーマットする必要はない。

【0027】図 4 は、本発明に関する範囲でのサーバ・オペレーティング・システム内でのファイル・サーバの動作を示し、図 5 は、ビデオ・データを求める要求に対するルータの DASD 割振り動作を示す図である。ファイル・サーバ・プログラムは、ステップ 201 で、ネットワークからのビデオ要求の処理と、ネットワークへのビデオ・データまたは接続失敗メッセージの配布を行う。次に、ステップ 203 で、ビデオ・ファイル・データに関する入力要求を、ルータに出力する。ステップ 205 で、ルータからビデオ・データまたは接続失敗メッセージを受け取る。その後、このプログラムはステップ 201 に戻る。

【0028】ルータ・プログラムは、ステップ 211 から始まり、ここで、ファイル・サーバによって転送されたビデオ接続要求を受け取る。接続要求を受け取って

ない場合、ステップ 223 に進んで、ビデオ・セグメントが完了したかどうかを判定することができる。ビデオ接続要求を受け取った場合、ステップ 213 で、駆動装置情報テーブルへのアクセスを行って、駆動装置使用可能性情報すなわち、駆動装置が走行中であるかどうかと、同時データ・ストリームの最大数未滿をサポートしているかどうかの情報を回復する。ステップ 215 で、駆動装置を使用できるかどうかを判定する。駆動装置を使用できる場合、ステップ 215 の YES 分岐からステップ 217 に進んで、使用可能な駆動装置のうちのどれを未処理の要求に割り当てるかを決定する。次に、ステップ 219 で、選択された DASD 装置で選択されたビデオを検索するのに適したコマンドを、割り当てられた DASD 装置の駆動装置制御装置に転送する。ステップ 221 で、駆動装置情報テーブルを更新する。最後に、ステップ 223 を実行して、前に要求されたビデオ・セグメントが完了したかどうかを判定する。そのようなビデオ・セグメントが完了している場合、ステップ 221 に戻って、駆動装置情報テーブルを更新する。完了したビデオ・セグメントがない場合、NO 分岐からステップ 211 に戻る。

【0029】ビデオ要求を受け取った時にすべての駆動装置が使用中になっている可能性がある。上で述べたように、ステップ 215 で、駆動装置が使用可能であるかどうかを判定している。使用可能な駆動装置がない場合、ステップ 215 の NO 分岐からステップ 225 に進んで、接続失敗表示をファイル・サーバに返す。

【0030】複数のミラーリングされた DASD 装置を管理するのに RAID 制御方式を利用することによって、複製データを有する複数の駆動装置がネットワーク・システム管理担当者によって直接管理される方式に対して、データ維持が単純化される。論理的には、1 つの更新が 1 つの駆動装置だけに適用される。

【0031】DASD 装置あたりのユーザ数の平衡化は、本発明の RAID ルータまたは RAID フィルタ (ADD) によって自動的に行われる。すべてのユーザが、サーバの観点からは同一の論理駆動装置を使用している。最大ユーザ数に関する当初の計画は、セットアップ時に、アレイに必要な駆動装置の数を決定するためにネットワーク・システム管理担当者によって行われる。追加の駆動装置が必要になった場合でも、簡単に追加できる。

【0032】まとめとして、本発明の構成に関して以下の事項を開示する。

【0033】(1) サーバによって制御される複数の直接アクセス記憶装置のそれぞれに、複製ビデオ・セグメントを記憶するステップと、再生プラットフォームからの要求に応答して、駆動装置情報テーブルからビデオ・セグメントの検索のために直接アクセス記憶装置を選択するステップと、選択された直接アクセス記憶装置に、

ビデオ・セグメントを検索するよう指令するステップと、検索に関する選択された直接アクセス記憶装置の使用を反映するため、駆動装置情報テーブルを更新するステップと、ネットワークを介して再生プラットフォームにビデオ・セグメントを伝送するステップとを含む、ネットワークを介してファイル・サーバに接続された複数の再生プラットフォームに、フル・モーション・ビデオを配布する方法。

(2) 前記選択するステップが、複数の直接アクセス記憶装置のそれぞれに関する使用可能性情報を検索するため、駆動装置情報テーブルにアクセスするステップと、1 つまたは複数の直接アクセス記憶装置が使用可能である場合に、ビデオ要求をサポートするため使用可能な直接アクセス記憶装置のうちの 1 つを割り当てるステップとを含むことを特徴とする、上記 (1) に記載の方法。

(3) 前記選択するステップが、さらに、使用可能な直接アクセス記憶装置がない場合に、再生プラットフォームに接続失敗を知らせるステップを含むことを特徴とする、上記 (2) に記載の方法。

(4) ビデオ・セグメントに関する再生プラットフォームからの要求に応答して、複数の直接アクセス記憶装置へのアクセスを制御するルータに要求を転送するステップをさらに含む、上記 (3) に記載の方法。

(5) 前記選択するステップが、ルータから実行されることを特徴とする、上記 (4) に記載の方法。

(6) 前記選択するステップが、装置ドライバから実行されることを特徴とする、上記 (3) に記載の方法。

(7) 必要データ速度を、拡張属性としてビデオ・ファイルに付加するステップをさらに含む、上記 (3) に記載の方法。

(8) 複数のビデオ・セグメントの再生が、同時にサポートされることを特徴とする、上記 (7) に記載の方法。

(9) 複数のビデオ・セグメントの再生が、複数の再生プラットフォームで発生することを特徴とする、上記 (8) に記載の方法。

(10) 複数の再生プラットフォームと、ファイル・サーバと、複数のプラットフォームとファイル・サーバとを相互接続するネットワークと、複数の直接アクセス記憶装置と、直接アクセス記憶装置のそれぞれの状況情報を提供する駆動装置情報テーブルと、複数の直接アクセス記憶装置のそれぞれに記録された、複数の複製ビデオ・セグメントと、再生プラットフォームからファイル・サーバへのビデオ・セグメントに関する要求に応答して、駆動装置情報テーブルへの参照によってビデオ・セグメントの検索のために直接アクセス記憶装置を選択する手段と、選択された直接アクセス記憶装置に、ビデオ・セグメントを検索するよう指令する手段と、検索に関する選択された直接アクセス記憶装置の使用を反映するため、駆動装置情報テーブルを更新する手段とを含む、

10

20

30

40

50

データ処理システム。

(11) 前記選択する手段が、複数の直接アクセス記憶装置のそれぞれに関する使用可能性情報を検索するため、駆動装置情報テーブルをアクセスする手段と、1つまたは複数の直接アクセス記憶装置が使用可能であることに応答して、ビデオ要求をサポートするため使用可能な直接アクセス記憶装置のうちの1つを割り当てる手段とを含むことを特徴とする、上記(10)に記載のデータ処理システム。

(12) 前記選択する手段が、さらに、使用可能な直接アクセス記憶装置がないことに応答して、再生プラットフォームに接続失敗を知らせる手段を含むことを特徴とする、上記(11)に記載のデータ処理システム。

(13) ビデオ・セグメントに関する再生プラットフォームからの要求に応答して、複数の直接アクセス記憶装置へのアクセスを制御するルータに要求を転送する手段をさらに含む、上記(12)に記載のデータ処理システム。

(14) 前記選択する手段が、ルータであることを特徴とする、上記(13)に記載のデータ処理システム。

(15) 前記選択する手段が、装置ドライバであることを特徴とする、上記(12)に記載のデータ処理システム。

(16) 必要データ速度を示す、選択されたビデオ・ファイルに対する拡張属性をさらに含む、上記(12)に記載のデータ処理システム。

(17) 複数のビデオ・セグメントの再生が、同時にサポートされることを特徴とする、上記(16)に記載のデータ処理システム。

(18) 複数のビデオ・セグメントの再生が、複数の再生プラットフォームで発生することを特徴とする、上記

(17)に記載のデータ処理システム。

\*

\*【0034】好ましい実施例に関して本発明を具体的に図示し、説明してきたが、当業者であれば、本発明の趣旨と範囲から逸脱せずに、形態と詳細にさまざまな変更を加えることができることを理解するであろう。

【0035】

【発明の効果】本発明によって、ドライブ名消費の問題が軽減され、完全なユーザ透明性がもたらされる。単一の駆動装置を、どのワークステーションからでもアクセスできる。

【図面の簡単な説明】

【図1】ネットワーク用ファイル・サーバに接続されたRAIDを含むデータ処理ネットワークの高水準ブロック図である。

【図2】複数の直接アクセス記憶装置上でRAIDを実施するようにプログラムされたファイル・サーバを含むデータ処理ネットワークの高水準ブロック図である。

【図3】図1のネットワークまたは図2のネットワークのいずれかで、ビデオ要求をサポートするためディスク駆動装置の割振りをサポートするのに使用されるデータ構造を示す図である。

【図4】本発明を実施するために修正されたファイル・サーバの動作の論理流れ図である。

【図5】本発明の実施に使用されるビデオ要求に関するルータ動作の論理流れ図である。

【符号の説明】

101 駆動装置IDフィールド

103 状況フラグ・フィールド

105 サポートされる最大ストリーム数フィールド

107 使用中フィールド

109 項目

139 緩衝記憶装置

【図3】

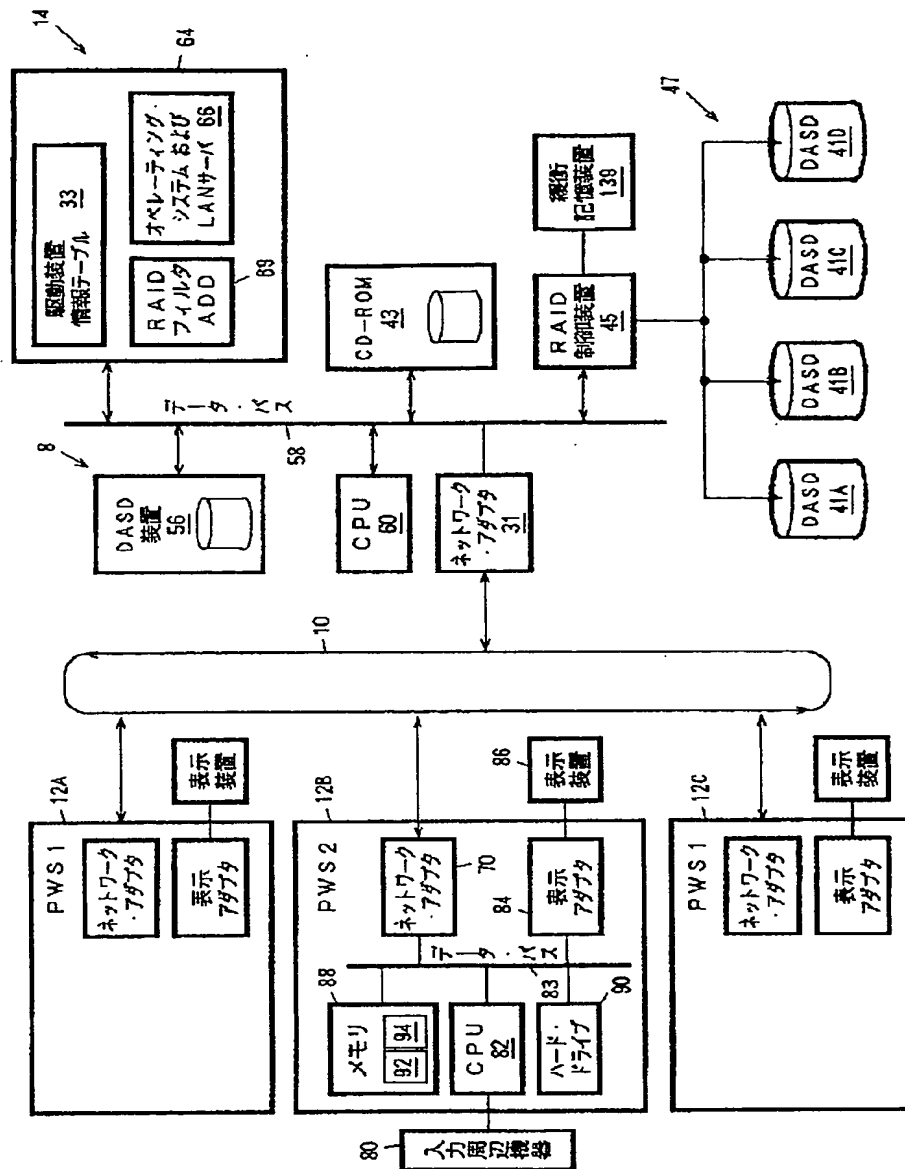
101 駆動装置 ID	103 状況 フラグ	105 サポートされる 最大ストリーム数	107 使用中	33
B1	1	5	2	108
B2	1	3	1	108
B3	1	4	2	109
B4	0	6	0	109

⋮

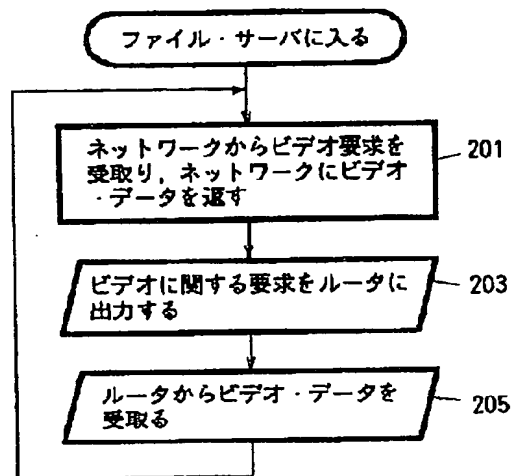


[illegible]

【図2】



【図4】



【図5】

